



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Gone Fishing

Citation for published version:

Bevan, S 2019, Gone Fishing: The Creation of the Comparative Agendas Project Master Codebook. in F Baumgartner, C Breunig & E Grossman (eds), *Comparative Policy Agendas: Theory, Tools, Data*. Oxford University Press, Oxford, United Kingdom, pp. 17-34. <https://doi.org/10.1093/oso/9780198835332.003.0002>

Digital Object Identifier (DOI):

[10.1093/oso/9780198835332.003.0002](https://doi.org/10.1093/oso/9780198835332.003.0002)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Comparative Policy Agendas

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Gone Fishing: The Creation of the Comparative Agendas Project Master Codebook

Shaun Bevan, University of Edinburgh, shaun.bevan@gmail.com

Abstract:

Every data gathering effort is a story, often a horror story from the perspective of those that created it. This chapter presents a historical tale of the creation and logic behind the Comparative Agendas Project (CAP) Master Codebook. The CAP is in reality a network of many projects aimed at classifying political agendas according to the policies they address. However, with no central administration or common source of funding the original coding framework experienced noticeable drift based on the context of each project. To harmonize the data across projects I led the creation of a common Master Codebook that was only possible with the support of the CAP community. This paper further discusses the limitations of the CAP data. Ultimately the master coded CAP data presents a common way of understanding policy attention and provides the framework for more detailed work in and outside the CAP community.

Keywords:

Master Codebook; Data Harmonization; Comparative Analyses; Coding; Validity

Acknowledgements:

Thanks go to the members of the CAP community for their hard work and hospitality. Particular thanks go to Christoffer Green-Pedersen, Stefaan Walgrave, Frank Baumgartner and Bryan Jones for their additional support and for originally suggesting and organizing this effort.

Special thanks go to the Mannheim Centre for European Social Research (MZES) at the University of Mannheim for allowing me to conduct this work as part of my research fellowship. Without their support the CAP Master Codebook would still only be just an idea.

For the past decade I have spent more time discussing fishing than I would have ever expected when I started my career in political science. Amazingly this is not because of an advisor's obsession with fishing,¹ the need to escape my work with a nice day along a river or due to the unending series of *Deadliest Catch* marathons on the Discovery Channel since I started my graduate training. Instead my time contemplating and deliberating on fishing has been about policy, namely the difficulties in conducting comparative analyses of public policy across nations. To allow the pun, this research note fishes through the creation of the Comparative Agendas Project (CAP) the finalization of the CAP Master Codebook and my time as a student, researcher, manager, faculty member and ultimately Director of the Master Codebook for the CAP. It is an effort to better understand CAP coding and data as well as the difficulties and limitations of a comparative approach to coding policy agendas.

Policy issues such as healthcare, national defense and social welfare are often are grouped in a logical manner based on one's own interpretation and understanding of the world. This inherent grouping extends to both human and computer coding techniques, including methods of scaling that group items based on the usage of keywords. However, as considerable work has shown scaling techniques are only applicable in a single language as different languages or even contexts can lead to considerable differences in the categories (e.g. Kluver 2009). As datasets, languages, countries and time periods change so does our interpretation and understanding of the world along with the usage and the meaning of words. In short, when a coding system is created, regardless of the means, it inherently matches the context it is created in. This background represents a massive challenge for the study of comparative policy agendas as no two contexts are ever exactly alike. As it turns out fisheries, policing, culture and many other policy areas present interesting challenges for comparative public policy as how these policies are defined and addressed varies considerably from nation

¹ My advisor is more of a tennis man actually.

to nation. To that end, this research note openly discusses the challenges of the CAP and introduces the CAP Master Codebook discussing its creation, intended use and limitations. The chapter is a guide for those interested in the ever-growing volume of CAP coded data that highlights the logic behind the CAP codes. It is my hope that this explication will lead to a better understanding of what can be done with CAP codes, how they can be iterated on and the wider use of CAP coded data as well as new projects that value the approach.

The rest of this chapter takes the following form. First, I discuss the creation of the CAP from its roots in the US Policy Agendas Project to its guiding principles of a limited coding system focused on coding policies, not targets. I next move onto the Master Codebook process explaining its necessity and the process of creating it. I further discuss some of the more difficult to address issues cross nationally and present a discussion of external validity. I conclude with an overview of the value and the limitations CAP data for current and future research. In addition, several appendices are tied to this chapter including a set of basic coding rules, the continuing Master Codebook process for new and existing projects and a brief introduction as well as guidelines for starting new projects. Up to date versions of these appendixes are maintained at [\[insert CAP link to book materials here\]](#).

The Comparative Agendas Project: A Philosophy and a Beginning

The CAP was built on the shoulders and limitations of the US Policy Agendas Project (US PAP) created by Frank Baumgartner and Bryan Jones. Widely used by political scientists, practitioners and students the US PAP data represents a key achievement for public policy research in the US. The project's goal was to create a series of commonly coded databases focused on the policy content of government and public agendas since World War II. To achieve this goal the US PAP codebook was created based on the development of the project's congressional hearings dataset with major and subtopic codes, the method for classifying data employed by the CAP, reflecting the policy attention of the US Congress. As

good planning, luck and/or serendipity would have it hearings are in fact a highly representative government agenda leading to the creation of a robust and lasting codebook for US government agendas (see Jones and Baumgartner 2005).

The codebook did however go through several revisions since the project was first started in 1993, such as the folding of family issues into law, order and family issues² and the development of several new subtopic codes in the major topic health. It has also been modified to fit new datasets. Even before the development of comparative projects challenges have routinely presented themselves such as how to address randomly sampled media data with stories on the weather, fires and obituaries presenting new topics not present in government datasets. Voluntary associations presented another problem with not all associations interested in policy such as Bob's International an association of people named Bob with awards for, among other things, the best Shiska-Bob (Bevan et al 2013). To deal

² The folding of immigration into labor and employment, and family issues into law and order is the reason for the missing major topic numbers 9 and 11 respectively in the original US codebook. The decision to combine these topics in the very early days of the project was based on how government addressed these issues in practice. For example, the US government purposefully avoids most family issues, but when it does it tends to address legal issues like child custody and illegal acts such as domestic violence. The advantage of the old, non-sequential coding system was that the missing major topics used to provide a quick indication of someone's familiarity with the codebook when asked how many major topic codes existed with the uninformed, modal answer being 21 with the correct answer being 19. Tragically, the CAP Master Codebook makes the old uninformed answer of 21 correct with the introduction of two new major topic codes removing the applicability of this informal heuristic.

with these issues new, non-policy codes were developed and introduced uniquely for the datasets that required them.

Each of these changes followed the “prime directive,” to quote Bryan Jones, of the CAP if you will which states that existing codes may never be combined, but that new codes can be created either to match truly new concerns or through the further separation of existing codes. For example, sports, specifically sports scores and news unrelated to the business aspect of sports, warranted a new code when the media data was first coded in the US. Moving beyond the US, immigration, an existing subtopic code in the original US codebook, warranted an extensive separation into a detailed major and subtopic structure in the majority of European nations where the policy area receives extensive attention in relation to the EU (see Guiraudon 2000). By following the “prime directive” these and other changes to the original codebook were in theory easily reverted in order to create a harmonized, common codebook for comparative analyses. In practice the process of harmonization was more involved and led to several revisions to the original codebook discussed in the Coding Through Compromise section.

But what is CAP Data?

Up until April 2016 with the launch of the CAP website CAP data has been defined quite broadly, namely as any dataset using a version of the CAP codebook in order to capture the policy attention of different government and public datasets based on their textual content. Policy attention meaning the substantive focus of the policy used, proposed or discussed for each observation. These observations can vary, from individual laws or news stories, for example, to aggregated measures such as with ‘most important problem’ type measures that capture the general policy attention of the public (e.g. Bevan and Jennings 2014). Regardless of the unit of analysis each observation is coded based on a common set of rules for that dataset aimed at capturing the primary policy focus of the observation. For example, coding

newspaper stories based on their introductory paragraph or secondary legislation based on their explanatory notes which summarize highly technical legislation. In general CAP datasets are coded over a long time frame and as comprehensively as possible, including all known bills or executive speeches over several decades for example. The datasets themselves include as much information as possible, including links and/or identifying information for each case as well as the text or other information used to code the case when such information is available and legally shareable. CAP data is therefore also as transparent, replicable and contestable as possible.

This broad definition of what CAP data has been refined to a new gold standard based on the criteria for data included on the common CAP website (see Appendix C). Namely, the ability for data to be matched to the CAP Master Codebook. This standard exists to ensure the CAP can live up to the first part of its acronym, comparative, well into the future.

Coding Limited by Design

Despite the scale of the data gathering across the CAP, the effort put into the Master Codebook and the unprecedented level of detail and access to raw data, the CAP is limited by design. The CAP community includes a wealth of researchers from political science, sociology, communication and computer science amongst other disciplines. That says nothing for the wealth of subfields represented and research questions being asked by the members of CAP projects. In short, while everyone involved in the CAP is concerned with attention in some manner or another, why and how they look at attention almost always differs. Building such a diverse and differently motivated group of scholars led to the limited, but robust coding system it employs. After all, attention is the sine qua non of policy-making as a change in framing, preferences and/or direction requires that a policy is first attended to. While this focus limits how CAP coded data can be directly used, the transparency and inherent replicability of the CAP datasets allows for more detailed work on the framing, preferences

and other factors beyond policy attention that make up each case. Whenever possible CAP datasets include both a means of linking to the original documents and importantly whatever text or information that was used to code each observation. This allows the users of the data to locate policies related to specific problems or countries, to build on or further refine the CAP coding by adding frames or more specific breakdowns of policy and importantly to challenge how a case or cases have been coded. Like any dataset, CAP data has errors, but by limiting the coding's focus to a common interest in attention and making the data as transparent as possible the quality of the data and its continued quality is as robust as the communities resources can allow for.

Policy not Targets? Terrorism and the Economic Crisis

Targets and policies are not the same. Whether or not solutions search for policy windows (Kingdon 1995) the policies aimed at targets often come from several places and often need to. Terrorism is one target for policy that often requires many different policies to address. Terrorism is a problem that can highlight issues with specific policy areas such as military intelligence, airline safety, immigration and a host of other areas. Moreover, the responsibility for directly addressing acts of terrorism can fall to the police, the military or a combination of both depending on the system of government and the source of terrorism itself. The CAP codes the policies that address the problem of terrorism according to their substantive policy focus rather than simply terrorism as a target for policy.

The CAP coding system's emphasis on the substantive focus of policy is perhaps its most common criticism (e.g. Dowding et al 2016). Clearly based on the description above a complete look at terrorism using CAP data would require additional work to identify the policies that were aimed at addressing the problem, but the same is true for other problems as well. How the CAP coding system addresses an economic crisis is another important example of the difference between targets and policies. The economic crisis that occurred in the late

naughts, and that has had continued effects for a number of years since, was a problem of banking, consumer confidence, unemployment and more. In short, it was a macroeconomic problem. However, the policy solutions to this macroeconomic problem did not just focus on changing interest rates, lowered taxes and other macroeconomic tools, but also focused on creating jobs, supporting new businesses and addressing social welfare issues in order to combat increasing unemployment numbers. In fact, the politics of austerity pushed by many nations meant that addressing the economic crisis included policies, however contentious, that touched on nearly every policy area government deals with from healthcare to public lands at least when it came to government spending. Ultimately, the economic crisis was a shock, a large shock that affected many policy areas that the government dealt with for a considerable time.

Problems no matter where they come from can lead to many different policies in many different areas with targets for policy such as terrorism or the economic crisis of the naughts driving new policies in the majority of policy areas. This is not a flaw of the CAP system of coding, but a choice to focus on policies and not targets. However, different targets as well as different problems, like countries and regions, can easily be identified through a search of the raw CAP data.

Coding Through Compromise: The CAP Master Codebook

Each project and often each dataset requires specific adaptations of the codebook to address observations and topics that do not exist in other contexts. More often than not these changes include the adaptation of existing codes to match the context of the project or dataset in question. However, with the number of projects having grown to nearly two dozen as of July 2014, the lack of a hierarchical CAP leadership and various levels of resources created a noticeable level of codebook drift. Much of this drift was of course necessary as projects needed to adapt the original American codebook to fit different contexts, while still keeping

the key goal of comparison in mind. While each project does an excellent job of coding and reconciliation with initial coding agreements ranging between 75% and 90% before cleaning, each of these activities were completed independently for each project. Only the determined focus and collegial nature of the CAP community led to generally comparable datasets that have already led to several noteworthy findings, such as the general effect of core issues on government attention (e.g. Jennings et al 2011) and truly general patterns of public policy (e.g. Jones et al 2009).

Nevertheless, these analyses were not without their flaws due to cross-national codebook incompatibilities. Some clear incompatibilities such as how immigration was included in the codebook were obvious. For the yet unknown differences robustness checks performed by the authors of this early research such as the jack-knifing of topics to search for any influential issues (e.g. Jennings et al 2011) means that while these codebook issues were not severe enough to change the inferences gained from these analyses future work without a truly comparative coding system might not be so lucky. As of July 2014 more than 450 subtopics existed across 15 projects with complete or draft crosswalks from an initial list of 225 subtopics. While most of these revised subtopics introduced minor alternations of existing codes in order to deal with minor differences between projects, a clear need for a common, comparative Master Codebook existed.

The process of developing the CAP Master Codebook started early on in the CAP's life and has had various members of the CAP community involved with Herschel F. Thomas III and Jeroen Joly having acted as previous heads laying the groundwork through an independent assessment of each team's national language codebook. This partial crosswalk was used as a basis for comparison in several comparative papers (e.g. Baumgartner et al. 2011), but a great deal of work was left to be done, especially on the subtopic level where many errors and incompatibilities were left to be addressed. The rest of this section describes the logic and history of the Master Codebook. Appendix B details the continuing Master

Codebook process as well as providing an example of how to apply the Master Codebook crosswalk to project coded data.

The Master Codebook Process

In the summer of 2012 I was informally elected by a group of project PIs to act as the Topic Coding Coordinator for the CAP with the goal of finalizing the process through a true Master Codebook. The objectives of this effort were outlined in a memo to all teams explaining the concept and the process of creating the CAP Master Codebook and were as follows: (1) Create a common Master Codebook that allows for accurate comparisons across all CAP datasets; (2) Minimize the overall amount of work by seeking a common middle ground between projects rather than asking any project to use a particular country's codebook; (3) Whenever possible avoid the need for recoding with appropriate aggregations.

Objective 1 was a clear cut goal to not exclude any project from the Master Codebook as the power and ideal of the CAP is its comparative nature. The second Objective to create a new Master Codebook was both practical and diplomatic in that my other work in the community did not drive the process. While I have and continue to work on the US and UK projects neither of their codebooks were appropriate as a basis for comparative Master Codebook. This is not just because the US is one of the only countries with a secondary mortgage market and the UK still has a monarchy either. In reality, like all project codebooks, they were adapted to a specific case and could not fit policy cross-nationally in general. Finally, considering the immense amount of work conducted by each team by the summer of 2012 Objectives 2 and 3 further proved essential with many teams low or lacking any additional resources in order to conduct this work. These objectives reduced the work load

through a least common denominator Master Codebook that maintained a balance of detail and feasibility splitting the distribution of work between projects.

With these objectives in mind and through the support of the Mannheim Centre of European Social Research (MZES) alongside the various projects I picked up where the previous efforts had left off by asking each team to create an English language codebook and arranging face to face meetings with projects over the next year. Prior to each meeting teams completed a common coding exercise on a selected set of UK Acts of Parliament intended to highlight common issues and difficult cases that was graded and discussed in detail in the meeting. Also prior to each meeting I read each English language codebook in its entirety to identify possible drift and new interpretations of codes to be discussed in the face to face meetings. Finally, before my first meeting I created a draft Master Codebook with 21 major topics and roughly 230 subtopics as a point of reference based on the previous Master Codebook efforts.

With a list of issues and the “graded” coding exercises in hand I visited each team to discuss their codebook and coding efforts in detail over the course of two days. Using my notes and the draft Master Codebook as a guide we discussed how to deal with any drift and inconsistencies either through the need for the project to recode the data or through the use of a crosswalk that often combined topics that were difficult to bound cross-nationally such as policing.³ I left each team with a brief list of issues for the team to directly address along with many notes concerning my draft Master Codebook and how to rebuild the crosswalk. After my last face to face meeting in the late spring of 2013 I cross-referenced these notes to build a second version of the Master Codebook with 21 major topics and 213 subtopics.

³ Policing primarily varies in its structure with national and local or subnational police operating in unique ways from country to country.

With a memo outlining the major differences between the Master Codebook and many national project codebooks and the draft Master Codebook sent to all teams I presented and sought comments at the largest coding meeting yet at the 2013 CAP conference in Antwerp, Belgium. Following a difficult discussion at the meeting and several revisions to the explanation of the changes and the process of recoding the data the first crosswalks between the Master Codebook and each project were completed and sent to national teams for further comments. This process led to several small changes concerning the crosswalks, but ultimately resulted in a final version of the Master Codebook with 21 major topic codes and 213 subtopic codes although significant revisions to the names of these topics were made to make them more generally applicable across projects. Crosswalks based on this final version of the Master Codebook continue to be produced, revised and proofed based on individual project feedback. When a proofed version of a project crosswalk is produced it is added to the Master Codebook crosswalk and the project adds the Master Codebook major and subtopic codes to their data.

Major topic codes and names for the CAP Master Codebook are presented in Table 1. A complete list of all subtopic codes is available at the CAP website along with an up to date version of Appendix B which outlines the process of matching project coded data to the CAP Master Codebook.

[insert Table 1 about here]

The Devil in the Details

Despite this rather straightforward, but intensive process for creating the Master Codebook the effort was far from easy. In order to create a truly comparable Master Codebook the devil was absolutely in the details with seemingly easy to understand issues like fisheries and culture creating some of the most intense debates possible concerning the

coding system. Policy, after all, differs based on context and with a variety of different political and temporal contexts to address a common understanding of policy across projects was a difficult task. As of July 2014 a total of more than 450 different subtopics existed across 15 projects with completed or draft crosswalks. The majority of these subtopics offered slight revisions in order to cover some unique aspect of the political system in question. Others were more unique highlighting the importance of specific religions in a countries policy-making or chose to split existing codes like freedom of speech and religion into its component parts. Some new codes like fishing and culture however had no true analog in the original US codebook and served as a source for debate since before I first started working with the CAP in 2007.

Fishing, a primary means for agribusiness in many European nations, fell in a mixture of the original US codes loosely tied to agriculture and was never an issue in landlocked countries like Switzerland. Comparatively, however, fishing is at least as important as ranching in the US or food safety in the United Kingdom.⁴ Based on the importance of the policy area, fishing was added as a new subtopic under agriculture in the Master Codebook.

No less important were the newly created immigration and culture major topic codes employed by a large number of projects. Immigration, while a common issue in many countries has often focused on civil rights or, in countries with seasonal and/or illegal workers, on labor issues. Yet, as most EU scholars would argue immigration is a policy area unto itself having played a major role in the creation of the EU (Guiraudon 2000) and as a continued source of debate and policy-making in many EU member states exemplifying its importance. Culture, namely the preservation and promotion of culture and language also plays an important role in many countries. Concerns over EU, US and other international

⁴ The United Kingdom's scandal concerning the presence of horsemeat in frozen food in early 2013 did little to ease the number of jokes told concerning British cuisine.

influences on nations like France and Italy has led to the production of a large volume of cultural policy. While countries like the US produce far less cultural policy, the importance of cultural policy in many systems is clear. The US's broad influence on culture internationally through its entertainment and business industries in fact adds to the external validity of culture as a policy area with little concern or need to maintain US culture compared to a strong focus by countries like France focused on the preservation of its culture and language. Due to the importance of culture and immigration cross-nationally they too were added to the Master Codebook as new major topic codes.

Ultimately the Master Codebook addresses each of these and many other seemingly minor issues by assuring that each team addresses the related policy area in the same manor in relation to the Master Codebook through each team's use of the Master Codebook crosswalk. By employing a common Master Codebook with a common and established way of coding the CAP data is internally valid cross-nationally.

A Cold War Mentality: Addressing the Country and Regional Subtopics

Created in the 90's in reference to the US policy since World War II the original US codebook like the policy it focused on maintained a Cold War mentality and view of the world when introducing country and regional subtopics. Intended as a category of last resort, the regional and country subtopic codes under international affairs were used on items with a broad, non-specific focus on a country or region that could not be coded elsewhere. The purpose of these codes is an extremely important one though, as they allow for a separation of foreign and domestic items especially in media and other similar agendas that are likely to mention other parts of the world without producing policy implications back at home.

However, the choice of which regions and countries to focus on reflected a transitory view of the world with subtopics like Soviet Union and Former Republics based on how countries and regions were discussed in the US Congress historically rather than based on

theory or geography. As a result, the country and especially the regional subtopic codes were used inconsistently between projects with no common agreement or rule on where to place countries like Turkey and Egypt. Moreover, many projects recognized this shortcoming and chose to forgo the original system altogether and instead introduced a dummy variable system that indicated the countries present in each item with the most extensive usage occurring in media agendas. This produced one of the biggest practical and potentially financial problems for the Master Codebook. Completing an additional level of coding for all projects that did not originally use the new dummy variable system was cost prohibitive, but the usage of the dated, Soviet era country and regional codes of yore would be wasteful. Similarly, the creation of a new system based on geography or current geopolitical standings would also be wasteful and eventually just as dated. Instead, a compromise to combine these historically dated subtopic codes into a single specific country or region subtopic code was decided upon. The exception to this rule is the code for the EU and Western Europe due to the inability for many projects to separate these two items. This process involved the combination of all regional and country subtopics in the countries that have them, and the crosswalk to the new code when the country dummy coding system indicated a focus on another nation without a focus on the project's own country. In other words, when the item was purely international affairs such as the election of a new foreign president it should be coded in this general specific country or region subtopic code. While the general specific region or country subtopic code is a loss of information from both sides, the transparency of the CAP data allows for much more directed and theory driven country and regional focuses based on a search of the data.

The Validity of Policy Differences

The process of creating the CAP Master Codebook focused on the comparability of policies and the terms/concepts that construct them. Across languages, time and various

institutional forms the system is designed so that policies governing everything from the angle of vehicle headlights to the legality of a certain election campaign receives the same major and sub-topic code regardless of the time, institutions or translations that need to take place from one data point to another. However, it is easy to forget that comparative research and a comparative design for research is about both similarity and differences. Many of the discussions and much of the feedback I received during the process of creating the Master Codebook concerned policy areas that received very little if any attention in a context or country. However, I saw these concerns as good, qualitative affirmations of validity. While the CAP Master Codebook had to be completely uniform, the applicability of CAP coding did not, in fact it should not be. For example, fishing is an important if not fundamental issue for certain countries like Denmark, but at best a very limited issue for a landlocked country such as Switzerland. Being a landlocked or an oceangoing nation does not mean potential policy areas differ, only that their applicability and level of use does. In order for CAP data to be valid representations of policy areas variation is essential.

This section considers the validity of policy differences between projects, intuitions and time periods. It makes use of the publically available tools on the CAP website as of January 2018 presented with a stable link to the original Figures and data in order to promote the free investigation and interpretation of these differences by readers and other scholars. The policy differences presented here are both in no particular order and based on no particular theory or world view. Instead, they simply represent some of the most common targets of the “we don’t” and “is not a policy here” comments I received while working on the CAP Master Codebook.

Defense Policy: The United States Vs Switzerland

In some ways the major topic defense was made for the United States. Not only has the US military been involved in a large number of military actions since World War II, but

spending on defense far outstrips every other CAP country.⁵ That spending creates many points for policy-making as well, from procurement procedures to bases and much, much more.

Switzerland on the other hand is quite different in this regard. Despite being a country with mandatory military service it is also a neutral country which has not taken a major military action since 1815 with spending generally less than 1/3 of that spent by the United States as measured as a percentage of GDP. Overall this leads to less of a need to attend to defense from a policy perspective than in the United States as demonstrated in Figure 1 which show the number of Reports/Bills and Legislation for both countries from 1978 to 2008.

[insert Figure 1 about here]

Clearly there is a vast disparity in legislative activity on defense between the United States and Switzerland. Nevertheless, not everything dataset follows this same pattern. A comparison of the front page of the *Neue Zürcher Zeitung* and the *New York Times Index* from 1995 to 2003 (see Figure 2) where data is currently available through CAP shows a noticeably higher level of attention to defense issues in the Swiss media owing to its more external and international media viewpoint.

[insert Figure 2 about here]

Culture: Something France has and the United Kingdom Does Not?

If someone was to overhear many of the discussions concerning culture within the CAP over the years it would seem as if half of policy scholars thought culture was merely a

⁵ Based on October 2016 data.

fictional concept. In reality much of the debate in the network over culture has been focused on whether or not cultural policy exists. Perhaps unsurprisingly several countries simply do not have cultural policies as the desire to promote their national language, protect cultural industries like film, theatre and more is not strong enough or central enough for the government to take notice. In others the importance of these sorts of items is strong enough to lead to government policies, sometimes very many policies. The CAP Master Codebook treats culture as a topic for policy, but one that admittedly is not attended to equally by all nations. In fact the differences between countries like the United Kingdom and France are so pronounced that they make the comparison for defense in Figures 1-2 look strong. Figure 3 shows the percentage of laws passed in the United Kingdom (1945-2012) and France (1979-2012) on cultural policy. A period of non-overlapping data for the UK was chosen to show that while cultural policy in the UK is rare, it did regularly receive attention for a time.

[insert Figure 3 about here]

The difference in the production of cultural policy in France and the United Kingdom is quite clear to see. For the overlapping period almost no cultural laws are made in the United Kingdom while as much as 7.4% of the laws passed by France are cultural in a year. While France shows a higher average compared to the United Kingdom outright it is noteworthy that the passage of cultural laws was at one time far more regular in the United Kingdom. However, the production of cultural laws all but died after Thatcher became Prime Minister, not only for her and her party, but for the administrations that came after showing a distinct and lasting impact on policy-making.

Universal Health... Policy Attention

Not all policy areas are created unequally. Perhaps one of the most surprising early comparisons born out of the emerging CAP network was the comparison of policy-making attention to health in Denmark and the United States (Green-Pedersen and Wilkerson 2006). Green-Pedersen and Wilkerson's (2006) work showed that despite fundamental differences in each country's approach to healthcare, a general rise in amount and complexity of legislative attention (bills, hearings, debates and questions) from the 1950s to the early 2000s occurred in both countries. Factors like new technologies and aging populations suggest that if a country is producing healthcare policy it must continue to attend to health. Figure 4 extends this work presenting the percentage of hearings, questions and Prime Minister's questions on health in the US, Denmark, Spain and UK from 1982 to 2002 (or for as long as data is available).

[insert Figure 4 about here]

While not as strikingly similar nor clearly trending as in the previous work, questions related to health in each of these countries have fairly steady levels of attention marked by peaks in activity. This both demonstrates the general importance of health across these four countries and likely much more broadly across most CAP countries, but also that national variation especially relating to higher patterns of attention do exist. In fact the same can be said for many of the major topic and subtopic comparisons made across countries, not just those for health. Given the generally similar concerns of most governments general patterns of attention across countries and datasets are to be expected though and adds to validity of CAP data.

Conclusion

Through the creation of the CAP Master Codebook and associated crosswalks CAP data allows for an unparalleled level of comparison on policy between nations. It is my sincere hope that the CAP data and the efforts the community have made at harmonizing it spurs a host of new comparative research never before possible. As the work from the CAP community has already shown several original insights can be gained through truly comparative cross national data such as the common law of punctuated budgets (Jones et al 2009) or the general effect of core issues on government agendas (Jennings et al 2011). The ever growing volume of CAP data is a resource not just for public policy scholars or even scholars in general, but students, practitioners, the media and elected officials alike if the success of the US PAP is any indication.

CAP coded data was designed as a tool for understanding policy attention, but can be used as a basis for so much more. While CAP data obviously has several limitations and cannot answer all policy questions, the framework and datasets are intended not just for analyses as they stand, but to be built upon. A great deal of work already demonstrates how CAP coded data provides a stepping stone for understanding framing (e.g. Boydston 2013) and more complex issues (e.g. Annesley et al 2015). While I certainly hope that the comparability of the CAP data breeds much more research into policy attention, it would be a tragedy if its design as a framework for understanding policy attention was not exploited in different ways. CAP data is at its core a database. It is a database organized by a common and comparative system for classifying policy attention that can be queried in order to locate observations of interest. Like a database of media stories it can be used to assess the level of attention, but can also be used for so much more with a more detailed and importantly directed investigation of the data. The CAP community agrees on the importance of attention in all of our research, but work on framing, preferences and more can and is being done based on this data. I think I am safe to say that as a community we only hope that the users of this database are able to add to and manipulate it in order to answer detailed and comparative

questions on policy in a way that was never before possible. If you are interested starting a project of your own and joining the CAP community please refer to the advice for new projects contained in Appendix C.

References

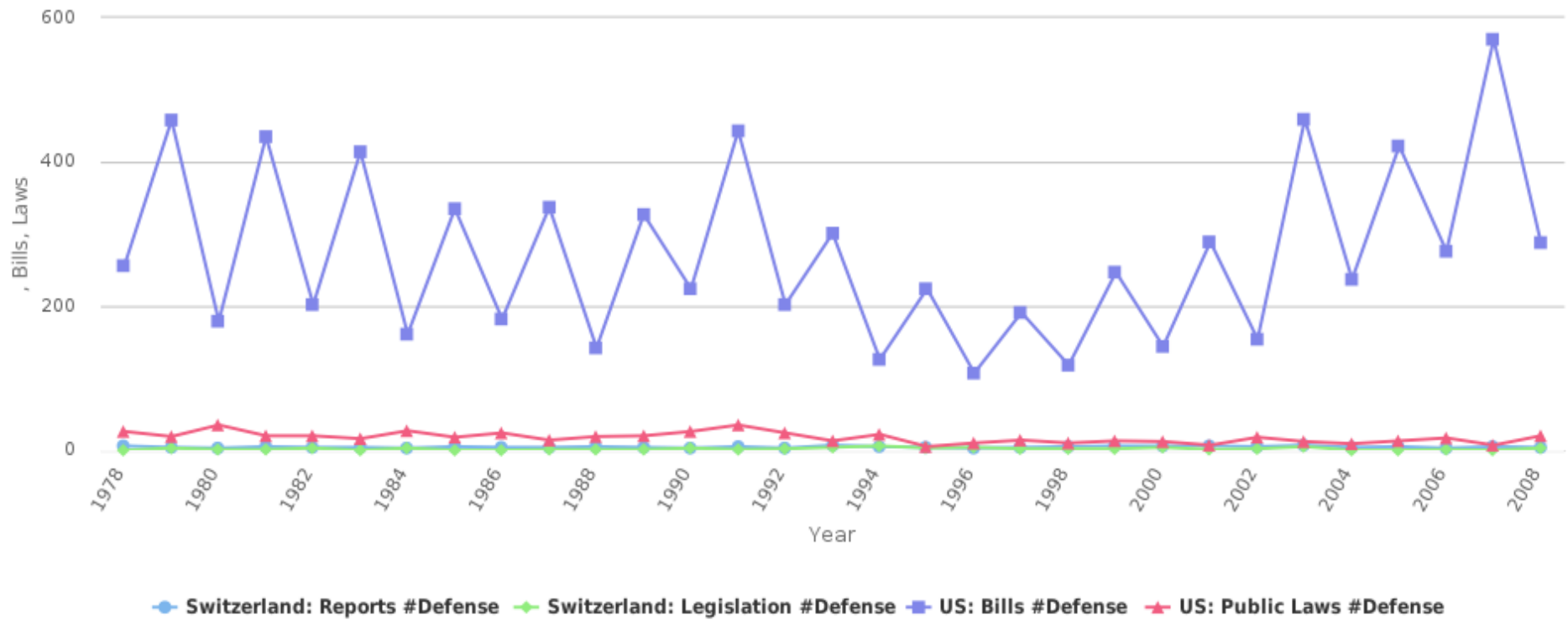
- Annesley, Claire, Isabelle Engeli, and Francesca Gains. 2015. "The profile of gender equality issue attention in Western Europe." *European Journal of Political Research* 54(3): 525-542.
- Baumgartner, Frank R., Sylvain Brouard, Christoffer Green-Pedersen, Bryan D. Jones and Stefaan Walgrave (Eds). 2011. "The Dynamics of Policy Change in Comparative Perspective" *Comparative Political Studies* 44(8).
- Bevan, Shaun, Frank Baumgartner, Erik Johnson and John McCarthy. 2013. "Understanding Selection Bias, Time-Lags and Measurement Bias in Secondary Data Sources: Putting the Encyclopedia of Associations Database in Broader Context." *Social Science Research*, 42(6): 1750-1764.
- Bevan, Shaun and Will Jennings. 2014. "Representation, Agendas and Institutions." *European Journal of Political Research*, 53(1): 37-56.
- Boydston, Amber E. 2013. *Making the News: Politics, the Media, and Agenda Setting*. Chicago: University of Chicago Press.
- Dowding, Keith, Andrew Hindmoor, and Aaron Martin. 2016. "The Comparative Policy Agendas Project: Theory, Measurement and Findings." *Journal of Public Policy* 36(1): 3-25.
- Green-Pedersen, Christoffer, and John Wilkerson. 2006. "How agenda-setting attributes shape politics: basic dilemmas, problem attention and health politics developments in Denmark and the US." *Journal of European Public Policy* 13(7): 1039-1052.
- Guiraudon, Virginie. 2000. "European Integration and Migration Policy: Vertical Policy Making as Venue Shopping." *Journal of Common Market Studies* 38 (2): 251–71.
- Jennings, Will, Shaun Bevan, Arco Timmermans, Gerard Breeman, Sylvain Brouard, Laura Chaques, Christoffer Green-Pedersen, Peter John, Anna Palau and Peter B. Mortensen.

2011. "Effects of the Core Functions of Government on the Diversity of Executive Agendas." *Comparative Political Studies*, 44(8): 1001-1030.
- Jones, Bryan D. and Baumgartner, Frank R. 2005. *The politics of attention: How government prioritizes problems*. Chicago: University of Chicago Press.
- Jones, Bryan D., Frank R. Baumgartner, Christian Breunig, Christopher Wlezien, Stuart Soroka, Martial Foucault, Abel François, Christoffer Green-Pedersen, Peter John, Chris Koski, Peter B. Mortensen, Frédéric Varone, and Stefaan Walgrave. 2009. "A general empirical law of public budgets: A comparative analysis." *American Journal of Political Science* 53(4) (2009): 855-873.
- Kingdon, John W. 1995. *Agendas, Alternatives, and Public Policies*. 2d. ed. New York: HarperCollins.
- Klüver, Heike. 2009. "Measuring interest group influence using quantitative text analysis." *European Union Politics* 10(4): 535-549.

Table 1: CAP Master Codebook Major Topic Codes

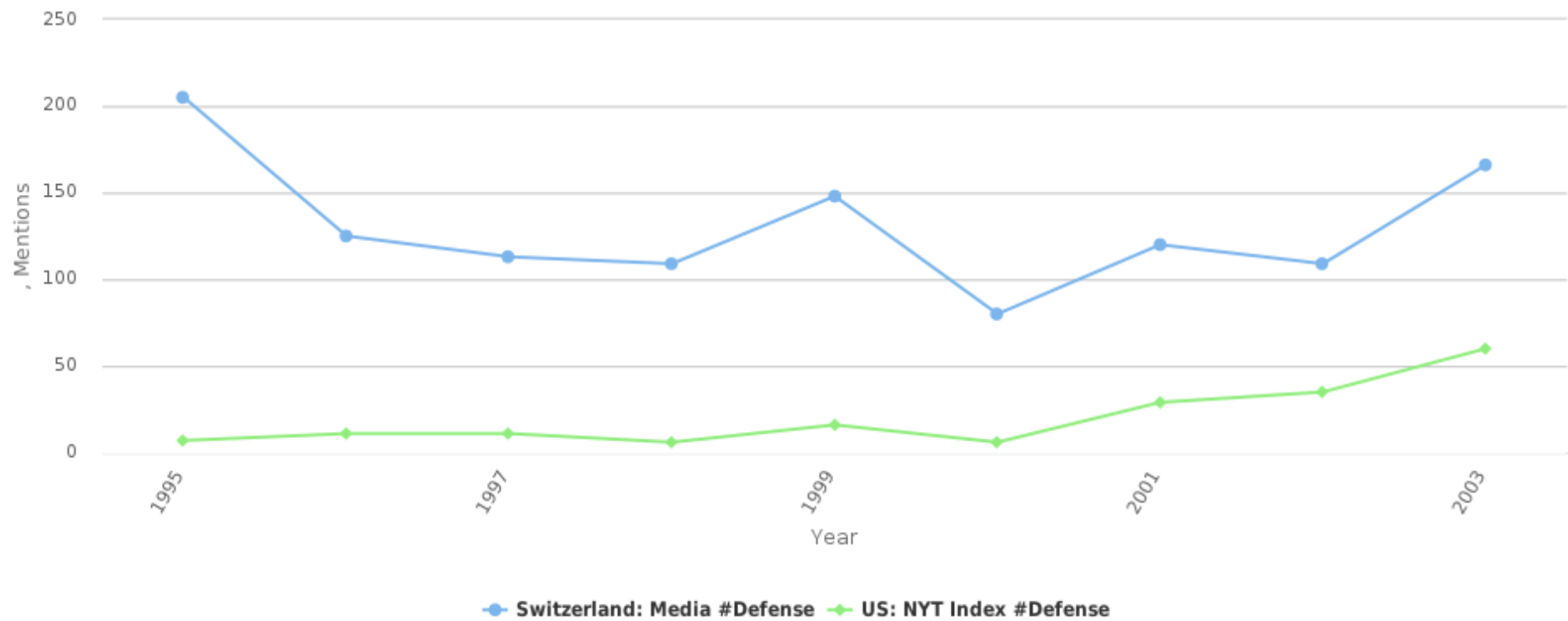
Major Topic	Title
1	Domestic Macroeconomic Issues
2	Civil Rights, Minority Issues, and Civil Liberties
3	Health
4	Agriculture
5	Labor and Employment
6	Education
7	Environment
8	Energy
9	Immigration and Refugee Issues
10	Transportation
12	Law, Crime, and Family Issues
13	Social Welfare
14	Community Development and Housing Issues
15	Banking, Finance, and Domestic Commerce
16	Defense
17	Space, Science, Technology, and Communications
18	Foreign Trade
19	International Affairs and Foreign Aid
20	Government Operations
21	Public Lands, Water Management, and Territorial Issues
23	Cultural Policy Issues

Figure 1: Bills and Laws – United States vs. Switzerland on Defense



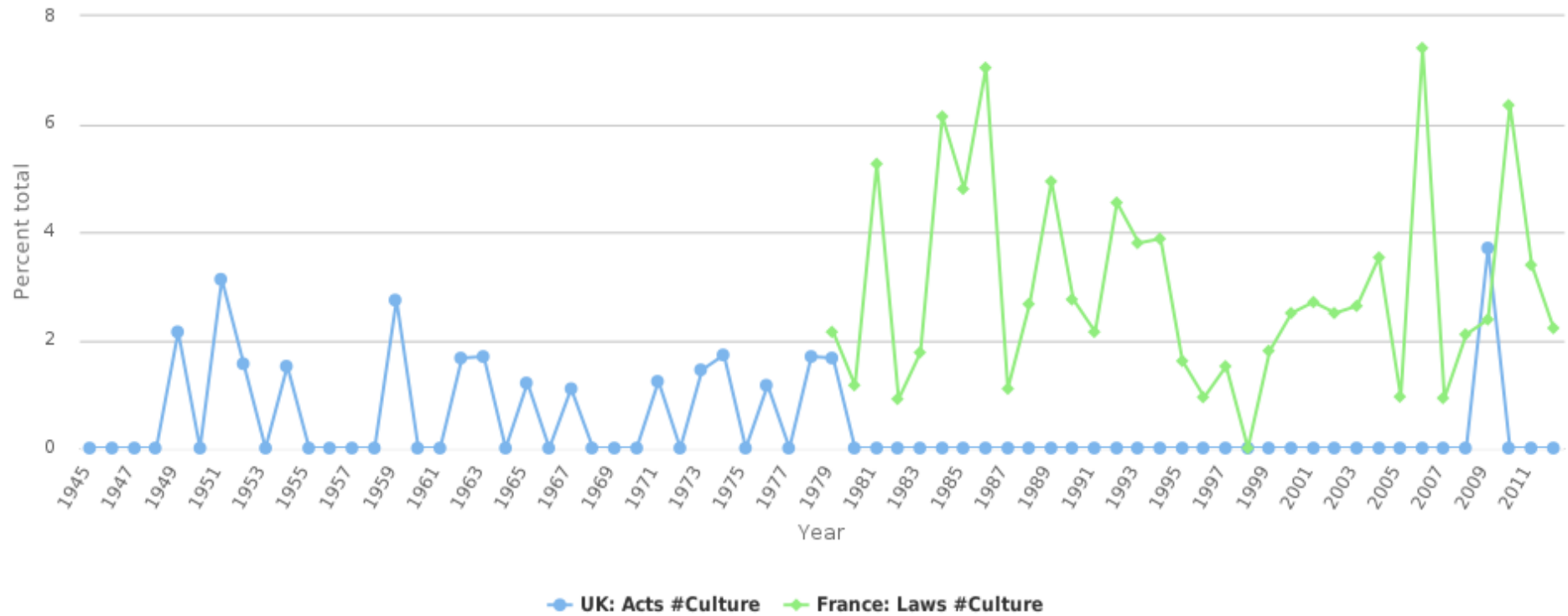
<http://www.comparativeagendas.net/tool/mAEptPK0>

Figure 2: Media - United States vs. Switzerland on Defense



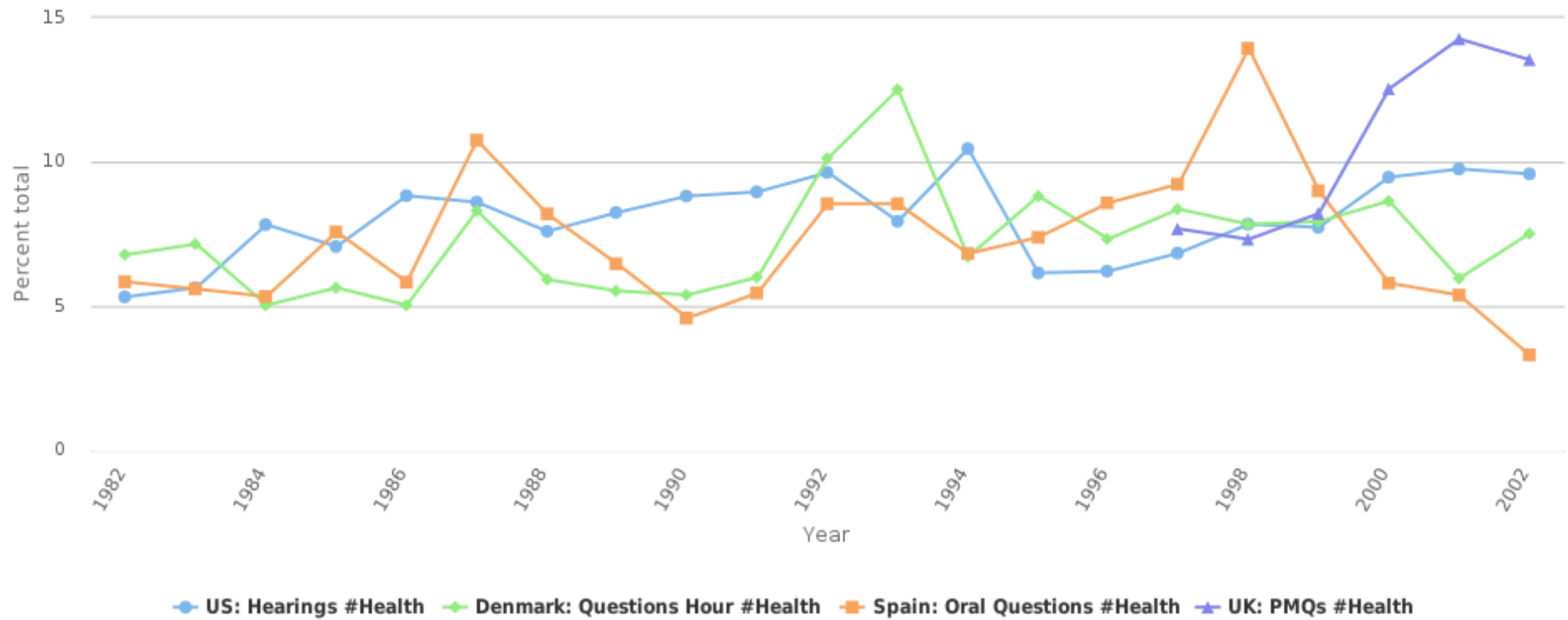
<http://www.comparativeagendas.net/tool/DOxSzImu>

Figure 3: Culture Across the Chanel – Cultural Laws in the United Kingdom and France



<http://www.comparativeagendas.net/tool/R155jHyZ>

Figure 4: Questioning Health – Legislative Questions Cross-Nationally



<http://www.comparativeagendas.net/tool/Ycb7NzXd>

Note: These Appendixes are not included in the final version of the chapter and instead reference other parts of the book or common CAP website. These Appendixes have been included as historical, Pre-Proof versions.

Appendix A

General Comparative Agendas Project Coding Guidelines

Created by: Shaun Bevan

Version: 0.9 Beta

Last Updated: 31 July 2014

See: sbevan.com/cap-master-codebook.html for an up to date version of this document

Note: These guidelines are adapted from the 2014 US Policy Agendas Project Codebook and follow from the general discussions that have occurred as part of the Master Codebook process.

1. Observations are coded according to the single predominant, substantive policy area rather than the targets of particular policies or the policy instrument utilized.
 - a. For example, if a case discusses changes to the home mortgage tax deduction, it would be coded according to the substantive policy area (consumer mortgages, code 1504) rather than the policy instrument (the tax code, code 107). In other words, while taxation is the tool being used in this example the policy area is consumer mortgages.
2. The “general” (00) subtopic includes cases where more than one distinct subtopic was discussed within a single major topic area.
 - a. For example, if a case discusses both drinking water (code 701) and air pollution (code 705), it would be coded as a general environmental issue (code 700). Thus, the general category within each major topic area includes some cases that are truly general as well as some cases that are the combination of as few as two subtopics. Each major topic includes an “other” category (NN99) for issues that do not fit into any of the categories and for which there were too few cases to justify the creation of a new category. For the Master Codebook this further applies to country specific subtopics, such a royal issues in the UK (2099 for the Master Codebook).
3. Observations that discuss appropriations for particular departments and agencies are coded according to their substantive policy area. Those that discuss appropriations for multiple departments and agencies that span multiple major topic codes are coded as general government operations (code 2000). While complete budgets and budget proposals are coded as national budget and debt (code 105).
 - a. For example, cases that discuss appropriations for the Ministry or Department of Energy are coded as energy (code 800) and those that discuss appropriations

railroads are coded as railroad transportation and safety (code 1005). Cases that discuss appropriations across multiple major topic areas, such as appropriations for diplomatic services (code 1900), national defense (code 1600), and the Ministry or Department of Energy (code 800), are coded as general government operations (code 2000). A proposed final budget or the total budget is coded as national budget and debt (code 105).

4. Observations that discuss terrorism and homeland security issues are coded according to their substantive policy focus.
 - a. For example, stricter airport security as a means for preventing terrorism should be coded as air transportation (code 1003). However, acts of terrorism and non-specific solutions should be coded as follows:
 - i. 1227: Policing and other domestic security responses to Terrorism (e.g. special police).
 - ii. 1615: Military and other national security responses to Terrorism (e.g. Homeland Security).
 - iii. 1927: International terrorism/hijacking (e.g. acts of piracy, terrorist incidents in foreign countries).
5. Observations that mention countries or regions not included in the project's national, subnational or supranational context are coded as follows.
 - a. Observations that discuss implications focused on the project's case that are related to the policy of a foreign country are coded according to the substantive policy area. For example, if a case discusses the implications of a country's economic markets on the project's own markets it is coded as securities and commodities regulation (code 1502).
 - i. Note: It is assumed that discussions of foreign policies within government agendas include implications for the project.
 - b. Observations that discuss the project's case and another country in a dyadic relationship are coded within foreign trade (18NN), foreign affairs (19NN), or defense (16NN) depending on the substantive policy focus. The project's foreign policy issues are coded under foreign affairs (19NN) and non-specific bilateral agreements are coded according to the country mentioned (code 1910 or 1921).
 - c. Observations not mentioning the project's case are coded according to their substantive policy area within international affairs (19NN). For example, if a case discusses violations of human rights in a specific country it is coded as human rights (code 1925). Remaining cases which do not mention the project's case are coded according to the region or specific country mentioned. These include cases such as the tax system of another country (code 1910 or 1921), or political developments in another country (code 1910 or 1921).
 - i. Note: Due to the assumption noted in 5a, the second part of this rule is predominately focused on public agendas such as the news media which often includes foreign policy stories that do not have direct implications for the project's case.

6. Observations that discuss indigenous affairs (code 2102), and capital city issues (2014) are placed in their corresponding subtopic codes regardless of substantive policy area. This exception is due to the special status of such communities with their activities not generally tied to national or project level politics. These are the only exceptions to the first guideline above.
 - a. For example, a case that discusses revisions to the healthcare system on a Native American reservation would be coded as indigenous affairs (code 2102) rather than comprehensive healthcare reform (code 301).
7. The administration, sovereignty, and other issues specific to the governing or status of dependencies and territories should be coded as 2105 rather than under government operations (major topic 20). However, in many cases the affairs of dependencies and territories should be coded under the appropriate subtopic depending on the relationship between the territory and the parent nation. For example, comprehensive healthcare reform in Scotland is coded as 301 rather than 2105 due to Scotland's integration into the UK government up until the point of devolution. How this rule was applied for each project should be well documented as the exact relationship between each nation and its dependencies varies greatly from country to country as does the level of national government involvement in the dependencies own affairs.

Note: Due to peculiarities in the data for some projects the so-called lowest code rule as an arbitrary means of making decisions has occasionally been used. The rule states:

While it is uncommon that observations not related to appropriations equally span two *major* topic areas, these observations are assigned the numerically lower major or subtopic code.

Through my various meetings the use of this rule is extremely limited and should only be used as a last resort.

Appendix B

The Comparative Agendas Project Master Codebook: Details and Process

Created by: Shaun Bevan

Version: 0.9 Beta

Last Updated: 31 July 2014

See: sbevan.com/cap-master-codebook.html for an up to date version of this document

This document outlines the continuing process of the Comparative Agendas Project (CAP) Master Codebook Crosswalk as of June 2014. It also includes a sample crosswalk procedure.

Please note that the Master Codebook crosswalk file is not annotated. Instead of the normal manner of annotation, this complete codebook consists of the cross-walks between each team's codebook and the Master Codebook (hence the repeated Master Codes). In this way those using the Master Codebook coded data will be able to reference each teams native language and/or English language codebooks providing the most accurate details for each dataset. While seemingly unwieldy, this file allows for a simple means of completing the crosswalk to the Master Codebook for each dataset as demonstrated below. As you can see the creation of new codes and splits of existing codes by individual projects as well as the new codes introduced through the Master Codebook itself led to 460 different subtopics across all projects currently contained in the Master Codebook crosswalk file. This further highlights the need for the Master Codebook for comparative work.

The Continuing Master Codebook Process:

1. Teams with a draft crosswalk that have not yet discussed this crosswalk with me should contact me to schedule a time discuss their crosswalk.
2. New teams that do not yet have a crosswalk should contact me to discuss the necessary resources I need to produce a crosswalk for their project.
3. Teams with a finalized crosswalk should proof this against the Master Codebook crosswalk file and please contact me with any errors or omissions so that these can be corrected.
4. Following the proofing of each finalized crosswalk against the Master Codebook crosswalk file I will produce an up to date version of the Master Codebook crosswalk to be placed on the common CAP data website. Teams should feel free to repost or link to this file however they see fit.

Completing the Crosswalk on your data:

The following example is based on Stata although the same general concept applies to other statistical software packages. If you create a version of this code for another software

package and would be willing to share it I would be happy to post it alongside the Master Codebook crosswalk file.

*** Change working directory (e.g. C:\Dropbox)

```
cd "C:\Dropbox"
```

*** Import the up to date crosswalk file

```
import excel "MasterCodebookCrosswalk.xls", sheet("Sheet1") firstrow
```

*** Save

```
save "MasterCodebookCrosswalk.dta", replace
```

*** Open an existing coded dataset (e.g. Codedlaws.dta)

```
use "CodedLaws.dta", clear
```

*** Rename your subtopic code to your country identifier (e.g. us for the United States)

```
rename subtopic us
```

*** Join the existing coded dataset to the saved MasterCodebookCrosswalk.dta by us

```
joinby us using "MasterCodebookCrosswalk.dta" update unmatched(both) _merge(Master)
```

*** Save

```
save "CodedLawsMasterCoded.dta", replace
```

This creates a version of your dataset including all original variables as well as the Master Codebook titles, major and subtopic codes.

Appendix C

The Comparative Agendas Project: Advice for New Projects

Created by: Shaun Bevan

Version: 0.9 Beta

Last Updated: 31 July 2014

See: sbevan.com/cap-master-codebook.html for an up to date version of this document

The Comparative Agendas Project Research Network

The Comparative Agendas Project (CAP) is a bit of a misnomer as the CAP is in reality a set of many projects aimed at classifying government and public agendas according to the policies they address. Functioning much like a network with many connections and key, central players there is no hierarchical leadership, no common budget and no formal rules for membership. Instead this most apolitical research community is focused on a common desire to understand policy attention based on a mixture of scientific disciplines, subfields, methods and questions.

Given the lack of formal institutions around the CAP research network depends on open and consistent communication with newly envisioned projects in order for the community to grow. While the CAP is not able to provide financial resources for the start-up of new projects, CAP principle investigators and researchers are happy to offer advice and feedback through every stage of a project from proposal, to coding and beyond. The CAP meeting held annually in June provides an excellent opportunity to meet members of the community, present a new project and to get continuous feedback on the wealth of work conducted by the community's members.

If you are interested in joining the community please feel free to e-mail us (comparativeagendas@gmail.com) to be added to the listserve and for advice on which team or teams might be the most helpful for the development of your project.

What to Study?

The CAP's focus on policy content attracts social scientists from a large number of backgrounds. While originally focused on replicating the US Policy Agendas Project in different nations through gathering the key executive and legislative agendas in each nation, projects have become increasingly diverse. National projects are still at the heart of the CAP, but subnational, supranational and multi-country projects often focused on a key agenda across existing or new CAP countries are also quite common in the community. In general, any government or public agenda that can be studied for its policy content fits well within the community. However, new projects will get the most value out of being part of the CAP community when they are able to draw on and be compared to existing CAP data. While the coding system has value on its own, the comparative nature of the CAP Master Codebook is clearly the biggest boon for comparative policy analyses.

Membership Criteria for the Common CAP Website

1. Your data must be able to be matched to the CAP Master Codebook on the major and/or subtopic level.

2. Project specific codebooks are allowed and generally encouraged as details that may not be comparable across projects may be very important for individual cases. However, these codebooks should follow Rule 1 above as well as the “prime directive,” that existing codes may never be combined, but that new codes can be created either to match truly new concerns or through the further separation of existing codes.
3. Data must be made freely available. While time limited embargos are understandable and sometimes even necessary to ensure the quality of the data, the ultimate output of the CAP is a set of large and freely available comparative datasets.
4. Data must be transparent. Whenever possible the original text or information used to code each item should be made available as well as the means of linking the data back to its original source.
5. A passion for understanding policy attention. While grant money and a large project team certainly help, many of the datasets produced by the community have been gathered by a single, motivated researcher without the aid of outside funds.

If your proposed project is able to fulfill the above membership criteria it is welcome to be hosted on the common CAP website. Please contact us for more details.